NOVEL REGULATORS OF FUNGAL GENE EXPRESSION

(Atty Docket No. MIC-004)

This is a continuation-in-part of provisional application Serial No. 60/257,431, filed 22 December 2000.

Field of the invention

The invention relates to regulators of fungal gene expression and their use in commercial and medical applications. More particularly, the invention relates to regulators of fungal genes involved in production of enzymes, secondary metabolites and other useful products, as well as to regulators of genes involved in fungal invasion.

Summary of the related art

Fungi are among the most common natural sources of useful substances for commercial and medical applications. Many of these useful substances are secondary metabolites.

Secondary metabolite production by various fungi has been an extremely important source of a variety of therapeutically significant pharmaceuticals. β-lactam antibacterials such as penicillin and cephalosporin are produced by *Penicillium chrysogenum* and *Acremonium chrysogenum*, respectively, and these compounds are by far the most frequently used antibacterials (reviewed in Luengo and Penalva, Prog. Ind. Microbiol. 29: 603-38 (1994); Jensen and Demain, Biotechnology 28: 239-68 (1995); Brakhage, Microbiol. Mol. Biol. Rev. 62: 547-85 (1998)). Cyclosporin A, a member of a class of cyclic undecapeptides, is produced by *Tolypocladium inflatum*. Cyclosporin A dramatically reduces morbidity and increases survival rates in transplant patients (Borel, Prog. Allergy 38: 9-18 (1986)). In addition, several fungal secondary metabolites are cholesterol-lowering drugs, including lovastatin that is made by *Aspergillus terreus* and several other fungi (Alberts *et al.*, Proc. Natl. Acad. Sci. USA 77: 3957-3961 (1980)). These and many other fungal secondary metabolites have contributed greatly to health

care throughout the world (see Demain, Ciba Found Symp 171: 3-16 (1992); Bentley, Crit. Rev. Biotechnol. 19: 1-40 (1999)).

Unfortunately, many challenges are encountered between the detection of a secondary metabolite activity and production of significant quantities of pure drug. Thus, efforts have been made to improve the production of secondary metabolites by fungi. Recently, strains have been improved by manipulating genes encoding the biosynthetic enzymes that catalyze the reactions required for production of secondary metabolites. Penalva *et al.*, Trends Biotechnol. 16: 483-489 (1998) discloses that production strains of *P. chrysogenum* have increased copy number of the penicillin synthesis structural genes. Other studies have modulated expression of other biosynthetic enzyme-encoding genes, thereby affecting overall metabolism in the fungus. Mingo *et al.*, J. Biol. Chem. 21: 14545-14550 (1999), demonstrate that disruption of the gene encoding phacA, an enzyme in *A. nidulans* that catalyzes phenylacetate 2-hydroxylation, leads to increased penicillin production, probably by elimination of competition for the substrate phenylacetate. Similarly, disruption of the gene encoding aminoadipate reductase in *P. chrysogenum* increased penicillin production, presumably by eliminating competition for the substrate alpha-aminoadipate (Casquiero *et al.*, J. Bacteriol. 181: 1181-1188 (1999)).

Thus, genetic manipulation holds promise for improving production of secondary metabolites. Genetic manipulation to increase the activity of biosynthetic enzymes for secondary metabolite production or to decrease the activity of competing biosynthetic pathways has proven effective for improving production. Maximum benefit might be achieved by combining several strategies of manipulation. For example, modulating the expression of genes that regulate the biosynthetic enzyme-encoding genes or altering concentrations of metabolic precursors might improve production. In addition, genetic manipulation could be used to impact upon the challenges that are encountered in the fermenter run or downstream processing (*e.g.* energy cost, specific production of desired metabolite, maximal recovery of metabolite, cost of processing waste from fermentations). There is, therefore, a need for regulator genes that can improve secondary metabolite production in a fungus.

Enzymes are another commercially important fungal product. Recently, efforts have been made to improve fungal enzyme production through genetic manipulation.

Noel *et al.*, Molecular Microbiology 27: 131-142 (1998), teaches that xlnR, a ZBC protein, induces expression of xylanolytic extracellular enzymes in *Aspergillus niger*. Hasper *et al.*, Molecular Microbiology 36: 193-200 (2000) teaches that xlnR also regulates D-xylose reductase gene expression in *Aspergillus niger*. Given the many useful enzymes produced by fungi, there remains a need for regulator genes that can improve the production of these enzymes.

There is also a need for regulator identifying genes relevant to fungal invasion. Fungal invasion is required for fungal pathogenesis and its regulation is likely related to secondary metabolite and enzyme production. Fungal infections have become a serious health concern, especially in immunocompromised patients. Ha and White, Antimicrobial Agents and Chemotherapy 43: 763-768 (1999) teach that candidiasis, which is caused by the pathogenic yeast *Candida albicans*, is the most frequent fungal infection associated with AIDS and other immunocompromised states. Weig *et al.*, Trends in Microbiology 6: 468-470 (1998) discloses that the frequency of Candida infections has increased in recent years and has been accompanied by a significant rise in morbidity and mortality. Many of these infections take place in the hospital setting. Baillie and Douglas, Methods in Enzymology 310: 644-656 (1999) teach that a majority of nosocomial septicemias caused by Candida species derive from biofilm formation on catheters and shunts. Little is known about the genes necessary for invasion or biofilm formation. There is, therefore, a need for the identification of new fungal invasion regulatory genes to act as targets for the development of antifungal drugs.

BRIEF SUMMARY OF THE INVENTION

The invention provides novel fungal regulator genes and methods for using regulator genes in commercial and medical applications.

In a first aspect, the invention provides novel isolated or recombinant genes that have been demonstrated to encode proteins that regulate fungal genes that are involved in secondary metabolite production, enzyme production, or fungal invasion. In certain preferred embodiments, the invention further provides homologs of such genes. These genes and their homologs are useful for improving secondary metabolite or enzyme production, or as targets for discovering new antifungal drugs.

In a second aspect, the invention provides isolated or recombinant nucleic acids that are specifically complementary to genes that have been demonstrated to encode proteins that regulate fungal genes that are involved in secondary metabolite production, enzyme production, or fungal invasion.

In a third aspect, the invention provides purified proteins that have been demonstrated to regulate fungal genes that are involved in secondary metabolite production, enzyme production, or fungal invasion. In certain embodiments, the invention further provides homologs of such proteins.

In a fourth aspect, the invention provides novel binding agents that specifically bind to proteins that have been demonstrated to regulate fungal genes that are involved in secondary metabolite production, enzyme production, or fungal invasion.

In a fifth aspect, the invention provides novel recombinant genes that are direct or indirect regulators of expression of *FLO11*, a fungal gene that is required for fungal invasion and whose expression is believed to be regulated by factors that also modulate secondary metabolite production. In certain embodiments, the invention further provides homologs of such genes. These genes are useful as targets for the development of antifungal drugs, and are expected to be useful for improving the production of secondary metabolites or fungal enzymes.

In a sixth aspect, the invention provides novel recombinant genes that are direct or indirect regulators of expression of *lovF*, a fungal gene involved in the production of the secondary metabolite lovastatin. In certain embodiments, the invention further provides

homologs of such genes. These genes are expected to be useful for improving the production of secondary metabolites or fungal enzymes.

In a seventh aspect, the invention provides novel recombinant genes that are direct or indirect regulators of expression of *lovE*, a fungal gene involved in the production of the secondary metabolite lovastatin. In certain embodiments, the invention further provides homologs of such genes. These genes are expected to be useful for improving the production of secondary metabolites or fungal enzymes.

In an eighth aspect, the invention provides novel recombinant genes that are direct or indirect regulators of expression of *acvA*, a fungal gene involved in the production of the secondary metabolite penicillin. In certain embodiments, the invention further provides homologs of such genes. These genes are expected to be useful for improving the production of secondary metabolites or fungal enzymes.

In a ninth aspect, the invention provides methods for modulating production of a secondary metabolite or enzyme, the method comprising expressing in the fungus a novel fungal regulator gene.

In a tenth aspect, the invention provides novel chimeric fungal regulator genes.

In an eleventh aspect, the invention provides methods for modulating production of a secondary metabolite or extracellular enzyme, the method comprising expressing in the fungus a novel chimeric fungal regulator gene.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The invention relates to regulators of fungal gene expression and their use in commercial and medical applications. More particularly, the invention relates to regulators of fungal genes involved in production of enzymes, secondary metabolites and other useful products, as well as to regulators of genes involved in fungal invasion. The patents and publications cited herein reflect the level of knowledge in this field and are hereby incorporated by reference in their entirety. In the case of conflict between the teaching of a cited reference and the present specification, the latter shall prevail.

The invention provides novel regulators of fungal gene expression, and methods for using regulator genes in commercial and medical applications.

In a first aspect, the invention provides novel isolated or recombinant genes that have been demonstrated to encode proteins that regulate fungal genes that are involved in secondary metabolite production, enzyme production, or fungal invasion. In certain preferred embodiments, the invention further provides homologs of such genes. These genes and their homologs are useful for improving secondary metabolite or enzyme production, or as targets for discovering new antifungal drugs.

As used herein, a "recombinant gene" is a nucleic acid sequence which encodes a protein, wherein the recombinant gene may be in the form of linear DNA or RNA, covalently closed circular DNA or RNA, or as part of a chromosome, provided however that it cannot be a single copy at the native chromosomal locus in which the gene normally resides in nature. Thus, this aspect of the invention includes genetically modified organisms containing such recombinant genes.

As used herein, a "homolog" of a gene is a gene that, when compared to the reference gene, has a lower E-value, or has a higher percent identity value in the region of homology than that listed in Table I. Table I shows a listing of reference genes, their E-values, and where applicable, their percent identity. In cases where regions of identity exist between the reference gene and a sequence from a public gene database, the region of identity is indicated. In such instances, position 1 is defined as the first nucleotide of the reference gene shown in the Sequence Listing.

In certain preferred embodiments, the homolog is a dominant mutation of the novel reference gene. A "dominant mutation" is an allele of a gene that encodes a protein capable of changing a phenotype of an organism more than a non-mutated form of the gene. Dominant mutations include, without limitation, mutations that encode a protein capable of changing a phenotype of an organism even when a non-mutant form of this gene (or its homologs) is resident in the organism. Preferred dominant mutations include dominant negative mutations, dominant positive mutations, and dominant neomorphic mutations. A "dominant negative mutation" is a dominant mutation that achieves its phenotypic effect by interfering with some function of the gene or gene product from which it was derived, or from a homolog thereof. A "dominant positive mutation" is a dominant mutation that achieves its phenotypic effect by activating some function of the gene or gene product from which it was derived, or from a homolog thereof. A "dominant neomorphic mutation" is a dominant mutation that achieves the phenotypic effect of providing a novel or different function to the gene or gene product from which it was derived, or from a homolog thereof.

Table I: Novel regulator genes

	Gene	E-score for coding sequence	Percent Identity (only for E<1e-25)	Starting Nucleotide	Ending Nucleotide
1	An10	6.00E-16	n/a		
2	An1000	1.00E-05	n/a		
3	At01-1	1.00E-05	n/a		
4	At01-2	7.00E-10	n/a		
5	At03	7.00E-12	n/a		
6	At05	1.00E-05	n/a		
7	At07	2.00E-06	n/a		
8	At08	1.00E-57	81	1	428
9	At11	8.00E-22	n/a		
10	At14	1.00E-05	n/a		
11	At16	7.00E-24	n/a		
12	At167	0	87	5	1514
13	At18	1.00E-05	n/a		
14	At19	1.00E-05	n/a		
15	At20	1.00E-05	n/a		
16	At22	3.00E-52	n/a		
17	At221	9.00E-17	n/a		
18	At233	9.00E-25	n/a		
19	At239	6.00E-30	84	763	939
20	At24	5.00E-07	n/a		
21	At240	7.00E-29	n/a		
22	At27	1.00E-05	n/a		

23	At274	1.00E-05	n/a		
24	At279	1.00E-05	n/a		
25	At286	1.00E-05	n/a		
26	At291	1.00E-05	n/a		
27	At32	1.00E-12	n/a		
28	At320	4.00E-15	n/a		
29	At322	1.00E-05	n/a		
30	At501	3.00E-21	n/a		
31	At574	2.00E-25	n/a		
32	Pc05	1.00E-05	n/a		
33	Pc06	1.00E-05	n/a		
34	Pc07	1.00E-05	n/a		
35	Pc08	1.00E-05	n/a		
36	Pc09	6.00E-16	n/a		
37	Pc10	1.00E-05	n/a		
38	Pc1000	2.00E-25	n/a		
39	Pc1001	1 00E-05	n/a		
40	Pc18	4.00E-12	n/a		
41	Pc23	1.00E-05	n/a		
42	Pc24	5.00E-18	n/a		
43	Pc25	1.00E-05	n/a		
44	Pc33	4.00E-07	n/a		
45	Pc34	6.00E-30	87	1042	1178
46	Pc804	1.00E-05	n/a		

Data shown in Table I are derived by comparing the novel reference genes with available genetic databases. The present inventors utilized the BLAST (Basic Local Alignment Search Tool) tool, which is publicly available at http://www.ncbi.nlm.nih.gov/blast/. This is a set of similarity search programs designed to explore sequence databases. The current version of NCBI BLAST, BLAST 2.1.2 was accessed directly from the NCBI web site. One of the outputs of the BLAST program is an Expect (E) value, which is a parameter that describes the number of "hits" one can expect to see by chance when searching a database of a particular size.

To determine E-values for the nucleotide coding sequences of the novel regulator genes according to the invention, the blastn program was used. This program compares a nucleotide query (reference) sequence against a nucleotide sequence database. The following databases available at the NCBI BLAST site were used: nr, which includes all GenBank, EMBL, DDBJ and PDB sequences; yeast (*Saccharomyces cerevisiae*) genomic nucleotide sequences; patents; and non-human/mouse EST (expressed sequence tag)

sequences. BLAST analysis was also performed using the *Neurospora crassa* partial genome sequence available at http://www.mips.biochem.mpg.de/cgi-bin/blast/blast_page?genus=neurospora. Table I reflects the lowest E-value for these five databases.

For a gapped alignment analysis, the blastn algorithm default settings were used. These are: -G, cost to open a gap, default = 5; -E, cost to extend a gap, default = 2; -q, penalty for a mismatch, default = -3; -r, reward for a match, default = 1; -e, expectation value, default = 10.0; and –W, word size, default = 11.

For novel reference genes having no related sequences in the databases, E-values were "floored" to 1e-5 as a way of defining a reasonable E-value for a hypothetical related gene. For genes having non-identical but related sequences in the databases, the E-value is that of the closest "relative" gene. For genes having extensive regions of homology to a gene in the database, a percent identity value provided by the algorithm is included in Table I, along with the region of identity.

Table II shows novel regulator genes according to the invention for which a portion of the gene sequence is in a database, but for which the present inventors now provide more complete sequence.

Table II: Novel regulator genes partially publicly disclosed

	Gene Name	Nucleotide Coding Sequence		Amino Acid Sequence		
		Microbia	Public	Microbia	Public	E score from BlastP of Microbia Proprietary AA
1	An01	506-649	1-505, 650-939	170-216	1-169, 217-313	1.00E-05
2	An05	1-1131	1132-1512	1-377	378-504	2.00E-13
3	An09	1-946	947-1557	1-315	316-519	1.00E-05
4	An13	532-823	1-531, 824-933	178-274	1-177, 275-311	1.00E-05
5	An17	1-457	458-660	1-152	153-220	1.00E-05
6	An20	28-90, 186-249	1-27, 91-185, 250-684	10-30, 62-83	1-9, 31-62, 84-228	1.00E-05
7	An28	521-572	1-520, 573-906	175-190	1-174, 191-302	1.00E-05
8	An34	1-119	120-532	1-39	40-178	1.00E-05
		533-810		179-270		1.00E-05

Preferred genes according to this aspect of the invention are selected from the group consisting of Pc804 (SEQ ID NO.1), An01 (SEQ ID NO. 3), An05 (SEQ ID NO.

5), An09 (SEQ ID NO 7), An10 (SEQ ID NO. 9), An13 (SEQ ID NO. 11), An17 (SEQ ID NO. 13), An20 (SEQ ID NO. 15), An28 (SEQ ID NO. 17), An34 (SEQ ID NO. 19), At01-1 (SEQ ID NO. 21), At01-2 (SEQ ID NO. 23), At03 (SEQ ID NO. 25), At05 (SEQ ID NO. 27), At07 (SEQ ID NO. 29), At08 (SEQ ID NO. 31), At11 (SEQ ID NO. 33), At14 (SEQ ID NO. 35), At16 (SEQ ID NO. 37), At18 (SEQ ID NO. 39), At19 (SEQ ID NO. 41), At20 (SEQ ID NO. 43), At22 (SEQ ID NO. 45), At24 (SEQ ID NO. 47), At27 (SEQ ID NO. 49), At32 (SEQ ID NO. 51), Pc05 (SEQ ID NO. 53), Pc06 (SEQ ID NO. 55), Pc07 (SEQ ID NO. 57), Pc08 (SEQ ID NO. 59), Pc09 (SEQ ID NO. 61), Pc10 (SEQ ID NO. 63), Pc18 (SEQ ID NO. 65), Pc24 (SEQ ID NO. 69), Pc25 (SEQ ID NO. 71), Pc33 (SEQ ID NO. 73), Pc34 (SEQ ID NO. 75), At501 (SEQ ID NO. 77), At574 (SEQ ID NO. 79), At279 (SEQ ID NO. 81), At286 (SEQ ID NO. 83), At291 (SEQ ID NO. 85), At320 (SEQ ID NO. 87), At322 (SEQ ID NO. 89), An1000 (SEQ ID NO. 91), At167 (SEQ ID NO. 93), At221 (SEQ ID NO. 95), At233 (SEQ ID NO. 97), At239 (SEQ ID NO. 99), At240 (SEQ ID NO. 101), At274 (SEQ ID NO. 103), Pc1000 (SEQ ID NO. 105), Pc1001 (SEQ ID NO.107).

In a second aspect, the invention provides isolated or recombinant nucleic acids that are specifically complementary to genes that have been demonstrated to encode proteins that regulate fungal genes that are involved in secondary metabolite production, enzyme production, or fungal invasion. The term "recombinant" is as used previously.

A sequence is "specifically complementary" to another sequence if it is sufficiently homologous to specifically hybridize to the sequence. A sequence "specifically hybridizes" to another sequence if it hybridizes to form Watson-Crick or Hoogsteen base pairs either *in vivo* or under conditions which approximate physiological conditions with respect to ionic strength, *e.g.*, 140 mM NaCl, 5 mM $MgCl_2$. Preferably, such specific hybridization is maintained under stringent conditions, *e.g.*, 0.2X SSC at 68°C.

In a third aspect, the invention provides purified proteins that have been demonstrated to regulate fungal genes that are involved in secondary metabolite production, enzyme production, or fungal invasion. In certain embodiments, the

invention further provides homologs of such proteins. As used herein, a "homolog" of a protein is a protein that, when compared to the reference protein, has a lower E-value or has a higher percent positive value in the region of similarity than that listed in Table III. Table III shows a listing of reference proteins, their E-values, and where applicable, their percent positives. In cases where regions of identity exist between the reference protein and a sequence from a protein sequence database, the positive region is indicated. In such instances, position 1 is defined as the first amino acid of the reference protein shown in the Sequence Listing. The percent positive score is defined by the BLAST 2.1.2 algorithm.

Table III: Novel regulator proteins

	Gene	E Score for Protein Sequence	Percent positive (only for E<1e-25)	Starting AA	Ending AA
1	An10	1.00E-166	85	7	372
2	An1000	1.00E-122	50	17	700
3	At01-1	1.00E-26	44	1	371
4	At01-2	1.00E-120	69	70	455
5	At03	2.00E-71	69	482	691
6	At05	4.00E-08	n/a		
7	At07	1.00E-11	n/a		
8	At08	0	78	1	430
9	At11	3.00E-95	46	488	1043
10	At14	5.00E-09	n/a		
11	At16	1.00E-10	n/a		
12	At167	0	97	1	510
13	At18	2.00E-08	n/a		
14	At19	1.00E-166	85	32	385
15	At20	2.00E-78	62	129	437
16	At22	0	84	1	506
17	At221	0	78	13	850
18	At233	6 00E-71	90	18	161
19	At239	1.00E-43	94	231	328
20	At24	3.00E-39	42	6	346
21	At240	0	82	1	571
22	At27	3.00E-24	n/a		
23	At274	1.00E-05	n/a		
24	At279	1.00E-95	55	1	496

25	At286	1 00E-156	81	71	412
26	At291	4.00E-10	n/a		
27	At32	0	72	4	796
28	At320	2.00E-47	76	1	126
29	At322	1.00E-05	n/a		
30	At501	1.00E-116	60	1	398
31	At574	1.00E-127	60	1	463
32	Pc05	8.00E-77	66	127	413
33	Pc06	7.00E-12	n/a		
34	Pc07	1.00E-05	n/a		
35	Pc08	2.00E-56	81	48	181
36	Pc09	4.00E-22	n/a		
37	Pc10	3.00E-31	68	76	189
38	Pc1000	1.00E-124	56	1	491
39	Pc1001	0	78	20	858
40	Pc18	7.00E-78	81	2	212
41	Pc23	3.00E-21	n/a		
42	Pc24	1 00E-10	n/a		
43	Pc25	5.00E-24	n/a		
44	Pc33	2.00E-72	38	43	852
45	Pc34	2.00E-65	87	241	395
46	Pc804	1.00E-05	n/a		

To obtain the values shown in Table III, a blastp alignment was performed against the respective nr, yeast and patent databases. This algorithm compares an amino acid query (reference) sequence against a protein sequence database. For this analysis, the following default settings were used for the blastp algorithm: -G, cost to open a gap, default = 11; -E, cost to extend a gap, default = 1; -e, expectation value, default = 10.0; and –W, word size, default = 3. In addition, the default substitution matrix, BLOSUM-62 was used to assign a score for aligning any possible pair of amino acid residues. The relevant settings for this matrix were gap existence cost = 11, per residue gap cost =1, and lambda ratio = 0.85. In cases where E-value was below 1e-25, a percent positive value is shown, along with the region over which this value occurs. In addition, coding sequence data is shown in Table II for proteins for which only a partial sequence has previously been known. Preferred proteins according to this aspect of the invention are selected from the group consisting of Pc804 (SEQ ID NO. 2), An01 (SEQ ID NO. 4), An05 (SEQ ID NO. 6), An09 (SEQ ID NO8), An10 (SEQ ID NO. 10), An13 (SEQ ID NO. 12), An17 (SEQ ID NO. 14), An20 (SEQ ID NO. 16), An28 (SEQ ID NO. 18), An34 (SEQ ID NO. 20), At01-1 (SEQ ID NO. 22), At01-2 (SEQ ID NO. 24), At03 (SEQ ID NO. 26), At05 (SEQ ID NO. 28), At07 (SEQ ID NO. 30), At08 (SEQ ID NO.

32), At11 (SEQ ID NO. 34), At14 (SEQ ID NO. 36), At16 (SEQ ID NO. 38), At18 (SEQ ID NO. 40), At19 (SEQ ID NO. 42), At20 (SEQ ID NO. 44), At22 (SEQ ID NO. 46), At24 (SEQ ID NO. 48), At27 (SEQ ID NO. 50), At32 (SEQ ID NO. 52), Pc05 (SEQ ID NO. 54), Pc06 (SEQ ID NO. 56), Pc07 (SEQ ID NO. 58), Pc08 (SEQ ID NO. 60), Pc09 (SEQ ID NO. 62), Pc10 (SEQ ID NO. 64), Pc18 (SEQ ID NO. 66), Pc23 (SEQ ID NO. 68), Pc24 (SEQ ID NO. 70), Pc25 (SEQ ID NO. 72), Pc33 (SEQ ID NO. 74), Pc34 (SEQ ID NO. 76), At501 (SEQ ID NO. 78), At574 (SEQ ID NO. 80), At279 (SEQ ID NO. 82), At286 (SEQ ID NO. 84), At291 (SEQ ID NO. 86), At320 (SEQ ID NO. 88), At322 (SEQ ID NO. 90), An1000 (SEQ ID NO. 92), At167 (SEQ ID NO. 94), At221 (SEQ ID NO. 96), At233 (SEQ ID NO. 98), At239 (SEQ ID NO. 100), At240 (SEQ ID NO. 102), At274 (SEQ ID NO. 104), Pc1000 (SEQ ID NO. 106), Pc1001 (SEQ ID NO.108).

"Purified", as used herein means having less than about 25% by weight, and preferably less than about 10% by weight contamination with other proteins. Such purified proteins may be obtained from natural sources, from recombinant expression, or by chemical synthesis. "Protein", as used herein is intended to encompass any polypeptide having at least 10 amino acid residues.

In a fourth aspect, the invention provides novel binding agents that specifically bind to proteins that have been demonstrated to regulate fungal genes that are involved in secondary metabolite production, enzyme production, or fungal invasion. Numerous methods familiar to those skilled in the art can be used to isolate novel binding agents. A binding agent "specifically binds" to a protein if it binds to that protein with greater affinity than to other unrelated proteins. Preferably, binding affinity of the molecule is at least 5-fold greater than affinity for unrelated proteins, more preferably at least 10-fold greater, still more preferably at least 50-fold greater, and most preferably at least 100-fold greater. In certain preferred embodiments, such binding agents decrease the biological function of a protein. In other preferred embodiments, binding agents increase the biological function of a protein. In other preferred embodiments, binding agents confer a novel or different biological function to a protein encoded by a fungal regulator gene. These effects may be achieved by increasing or decreasing protein activity through changes in secondary or tertiary structure, creating a new protein activity through

changes in secondary or tertiary structure; increasing or decreasing transcription, increasing or decreasing translation, increasing or decreasing post-translational modification, altering intracellular localization, increasing or decreasing translocation from one cellular location to another, increasing or decreasing protein activity by interaction of the protein with another molecule, or creating a new protein activity by interaction of the protein with another molecule.

In a fifth aspect, the invention provides novel isolated or recombinant genes and their encoded proteins that are direct or indirect regulators of expression of *FLO11*, a fungal gene that is required for fungal invasion and whose expression is believed to be regulated by factors that also modulate secondary metabolite production. In certain embodiments, the invention further provides homologs of such genes or proteins. The term "homolog" is as defined previously. These genes are useful as targets for the development of antifungal drugs, and are expected to be useful for improving the production of secondary metabolites or fungal enzymes. They are also useful as tools to study the role of *FLO11* in fungal metabolism.

Preferred genes according to this aspect of the invention are selected from the group consisting of An01 (SEQ ID NO. 3), An05 (SEQ ID NO. 5), An09 (SEQ ID NO 7), An10 (SEQ ID NO. 9), An13 (SEQ ID NO. 11), An17 (SEQ ID NO. 13), An20 (SEQ ID NO. 15), An28 (SEQ ID NO. 17), An34 (SEQ ID NO. 19), At01-1 (SEQ ID NO. 21), At01-2 (SEQ ID NO. 23), At03 (SEQ ID NO. 25), At05 (SEQ ID NO. 27), At07 (SEQ ID NO. 29), At08 (SEQ ID NO. 31), At11 (SEQ ID NO. 33), At14 (SEQ ID NO. 35), At16 (SEQ ID NO. 37), At18 (SEQ ID NO. 39), At19 (SEQ ID NO. 41), At20 (SEQ ID NO. 43), At22 (SEQ ID NO. 45), At24 (SEQ ID NO. 47), At27 (SEQ ID NO. 49), At32 (SEQ ID NO. 51), Pc05 (SEQ ID NO. 53), Pc06 (SEQ ID NO. 55), Pc07 (SEQ ID NO. 57), Pc08 (SEQ ID NO. 59), Pc09 (SEQ ID NO. 61), Pc10 (SEQ ID NO. 63), Pc18 (SEQ ID NO. 65), Pc23 (SEQ ID NO. 67), Pc24 (SEQ ID NO. 69), Pc25 (SEQ ID NO. 71), Pc33 (SEQ ID NO. 73), Pc34 (SEQ ID NO. 75).

Preferred proteins according to this aspect of the invention are selected from the group consisting of An01 (SEQ ID NO. 4), An05 (SEQ ID NO. 6), An09 (SEQ ID NO. 8), An10 (SEQ ID NO. 10), An13 (SEQ ID NO. 12), An17 (SEQ ID NO. 14), An20

(SEQ ID NO. 16), An28 (SEQ ID NO. 18), An34 (SEQ ID NO. 20), At01-1 (SEQ ID NO. 22), At01-2 (SEQ ID NO. 24), At03 (SEQ ID NO. 26), At05 (SEQ ID NO. 28), At07 (SEQ ID NO. 30), At08 (SEQ ID NO. 32), At11 (SEQ ID NO. 34), At14 (SEQ ID NO. 36), At16 (SEQ ID NO. 38), At18 (SEQ ID NO. 40), At19 (SEQ ID NO. 42), At20 (SEQ ID NO. 44), At22 (SEQ ID NO. 46), At24 (SEQ ID NO. 48), At27 (SEQ ID NO. 50), At32 (SEQ ID NO. 52), Pc05 (SEQ ID NO. 54), Pc06 (SEQ ID NO. 56), Pc07 (SEQ ID NO. 58), Pc08 (SEQ ID NO. 60), Pc09 (SEQ ID NO. 62), Pc10 (SEQ ID NO. 64), Pc18 (SEQ ID NO. 66), Pc23 (SEQ ID NO. 68), Pc24 (SEQ ID NO. 70), Pc25 (SEQ ID NO. 72), Pc33 (SEQ ID NO. 74), and Pc34 (SEQ ID NO. 76).

In a sixth aspect, the invention provides novel recombinant genes and their encoded proteins that are direct or indirect regulators of expression of *lovF*, a fungal gene believed to be involved in the production of the secondary metabolite lovastatin. In certain embodiments, the invention further provides homologs of such genes or proteins. The term "homolog" is as defined previously. These genes are expected to be useful for improving the production of secondary metabolites or fungal enzymes. They may also be useful as tools to study the role of *lovF* in fungal metabolism.

Preferred genes according to this aspect of the invention are selected from the group consisting of At279 (SEQ ID NO. 81), At286 (SEQ ID NO. 83), At291 (SEQ ID NO. 85), At320 (SEQ ID NO. 87), At322 (SEQ ID NO. 89), An1000 (SEQ ID NO. 91), At167 (SEQ ID NO. 93), At221 (SEQ ID NO. 95), At233 (SEQ ID NO. 97), At239 (SEQ ID NO. 99), At240 (SEQ ID NO. 101), At274 (SEQ ID NO. 103), Pc1000 (SEQ ID NO. 105), Pc1001 (SEQ ID NO.107).

Preferred proteins according to this aspect of the invention are selected from the group consisting of At279 (SEQ ID NO. 82), At286 (SEQ ID NO. 84), At291 (SEQ ID NO. 86), At320 (SEQ ID NO. 88), At322 (SEQ ID NO. 90), An1000 (SEQ ID NO. 92), At167 (SEQ ID NO. 94), At221 (SEQ ID NO. 96), At233 (SEQ ID NO. 98), At239 (SEQ ID NO. 100), At240 (SEQ ID NO. 102), At274 (SEQ ID NO. 104), Pc1000 (SEQ ID NO. 106), and Pc1001 (SEQ ID NO.108).

In a seventh aspect, the invention provides novel recombinant genes and their encoded proteins that are direct or indirect regulators of expression of *lovE*, a fungal gene believed to be involved in the production of the secondary metabolite lovastatin. In certain embodiments, the invention further provides homologs of such genes or proteins. The term "homolog" is as defined previously. These genes are expected to be useful for improving the production of secondary metabolites or fungal enzymes. They may also be useful as tools to study the role of *lovE* in fungal metabolism.

Preferred genes according to this aspect of the invention are selected from the group consisting of At501 (SEQ ID NO. 77), At574 (SEQ ID NO. 79).

Preferred proteins according to this aspect of the invention are selected from the group consisting of At501 (SEQ ID NO. 78), At574 (SEQ ID NO. 80).

In an eighth aspect, the invention provides novel recombinant genes and their encoded proteins that are direct or indirect regulators of expression of *acvA*, a fungal gene involved in the production of the secondary metabolite penicillin. In certain embodiments, the invention further provides homologs of such genes. The term "homolog" is as defined previously. These genes are expected to be useful for improving the production of secondary metabolites or fungal enzymes. They may also be useful as tools to study the role of *acvA* in fungal metabolism.

A preferred gene according to this aspect of the invention is Pc804 (SEQ ID NO. 1). A preferred protein according to this aspect of the invention is Pc804 (SEQ ID NO. 2).

In a ninth aspect, the invention provides methods for modulating production of a secondary metabolite or extracellular enzyme, the method comprising expressing in the fungus a novel fungal regulator gene. Such expression may be achieved and/or modulated by expression of another gene, interaction with a drug or small molecule, or genetic transformation, *e.g.*, as described in the Examples. In addition it may include expression of an entire gene or of a biologically active portion or fragment of the gene, or a fusion product of the gene or a portion or fragment of the gene. In certain preferred embodiments, combinations of more than one novel regulator gene according to the

invention may be expressed simultaneously or sequentially.

As used herein, the term "modulatinging production of a secondary metabolite or extracellular enzyme" means to positively impact upon one or more of the variables that affect the process of production of secondary metabolites or extracellular enzyme in a fungal fermentation. These variables include, without limitation, amount of secondary metabolite or extracellular enzyme being produced (in absolute terms or in quantity per unit volume of fermentation broth or per unit mass of solid substrate), the volume required for production of sufficient quantities, the cost of raw materials and energy, the time of fermenter or culture run, the amount of waste that must be processed after a fermenter run, the specific production of the desired metabolite or extracellular enzyme (both in total amounts and as a fraction of all secondary metabolites and side products made by the fungus), the percent of produced final secondary metabolite or extracellular enzyme that can be recovered from the fermentation broth or culture, and the resistance of an organism producing a secondary metabolite or extracellular enzyme to possible deleterious effects of contact with the secondary metabolite or extracellular enzyme. Also, the term "secondary metabolite" means a compound, derived from primary metabolites, that is produced by an organism, is not a primary metabolite, is not ethanol or a fusel alcohol, and is not required for growth under standard conditions. Secondary metabolites are derived from intermediates of many pathways of primary metabolism. These pathways include, without limitation, pathways for biosynthesis of amino acids, the shikimic acid pathway for biosynthesis of aromatic amino acids, the polyketide biosynthetic pathway from acetyl coenzyme A (CoA), the mevalonic acid pathway from acetyl CoA, and pathways for biosynthesis of polysaccharides and peptidopolysaccharides. Collectively, secondary metabolism involves all primary pathways of carbon metabolism (Fungal Physiology, Chapter 9 pp 246-274 ed DH Griffin (1994)). "Secondary metabolites" also include intermediate compounds in the biosynthetic pathway for a secondary metabolite that are dedicated to the pathway for synthesis of the secondary metabolite. "Dedicated to the pathway for synthesis of the secondary metabolite" means that once the intermediate is synthesized by the cell, the cell will not convert the intermediate to a primary metabolite. "Intermediate compounds" also include secondary metabolite intermediate compounds which can be converted to

useful compounds by subsequent chemical conversion or subsequent biotransformation. As such, providing improved availability of such intermediate compounds would still lead to improved production of the ultimate useful compound, which itself may be referred to herein as a secondary metabolite. The yeast *Saccharomyces cerevisiae* is not known to produce secondary metabolites. The term "primary metabolite" means a natural product that has an obvious role in the functioning of almost all organisms. Primary metabolites include, without limitation, compounds involved in the biosynthesis of lipids, carbohydrates, proteins, and nucleic acids. The term "increasing the yield of the secondary metabolite" means increasing the quantity of the secondary metabolite present in the total fermentation broth per unit volume of fermentation broth or unit mass of solid media. An "extracellular enzyme" is an enzyme that is secreted by a fungus into its surrounding media.

In a tenth aspect, the invention provides novel chimeric fungal regulator genes. For purposes of the invention, a "chimeric fungal regulator gene" is a fungal regulator gene, or a portion thereof, fused to a nucleic acid sequence to which it is not fused in nature. Preferred nucleic acids for such fusion purposes include, without limitation, fungal regulator genes, or portions thereof, as well as regulatory regions from a gene.

A non-limiting example of some preferred chimeric fungal regulator genes is shown in Table 4 below, and in the Sequence listing as SEQ ID NOS.

Table IV: Novel chimeric regulator proteins

	Chimera	species A*	species B**	n-terminal fusion (nt)	filler (nt)	c-terminal fusion (nt)
1	amdA-pacC(DBD)	*A. nidulans*	*P. chrysogenum*	1-468	469-504	505-1074
2	VP16-An09	Herpes simplex virus	*A. nidulans*	1-243	244-276	277-1833
3	VP16-Pc23	Herpes simplex virus	*P. chrysogenum*	1-243	244-276	277-1353
4	amdAG229C-pacC(DBD)	*A. nidulans*	*P. chrysogenum*	1-468	469-504	505-1074
5	amdAG229D-pacC(DBD)	*A. nidulans*	*P. chrysogenum*	1-468	469-504	505-1074
6	pacC(DBD)-VP16	*P. chrysogenum*	Herpes simplex virus	1-567	568-573	574-810
7	leu4-tet	*S. cerevisiae*	*E. coli*	1-1515	n/a	1516-2415

In an eleventh aspect, the invention provides methods for modulating production of a secondary metabolite or extracellular enzyme, the method comprising expressing in the fungus a novel chimeric fungal regulator gene. Such expression may be achieved and/or modulated by expression of another gene, interaction with a drug or small molecule, or genetic transformation, *e.g.*, as described in the Examples. In certain preferred embodiments, combinations of more than one novel regulator gene and/or chimeric reggene according to the invention may be expressed simultaneously or sequentially.

As used herein, the term "modulating production of a secondary metabolite or extracellular enzyme" means to positively impact upon one or more of the variables that affect the process of production of secondary metabolites or extracellular enzyme in a fungal fermentation. These variables include, without limitation, amount of secondary metabolite or extracellular enzyme being produced (in absolute terms or in quantity per unit volume of fermentation broth or per unit mass of solid substrate), the volume required for production of sufficient quantities, the cost of raw materials and energy, the time of fermenter or culture run, the amount of waste that must be processed after a fermenter run, the specific production of the desired metabolite or extracellular enzyme (both in total amounts and as a fraction of all secondary metabolites and side products made by the fungus), the percent of produced final secondary metabolite or extracellular enzyme that can be recovered from the fermentation broth or culture, and the resistance of an organism producing a secondary metabolite or extracellular enzyme to possible deleterious effects of contact with the secondary metabolite or extracellular enzyme. Also, the term "secondary metabolite" means a compound, derived from primary metabolites, that is produced by an organism, is not a primary metabolite, is not ethanol or a fusel alcohol, and is not required for growth under standard conditions. Secondary metabolites are derived from intermediates of many pathways of primary metabolism. These pathways include, without limitation, pathways for biosynthesis of amino acids, the shikimic acid pathway for biosynthesis of aromatic amino acids, the polyketide

biosynthetic pathway from acetyl coenzyme A (CoA), the mevalonic acid pathway from acetyl CoA, and pathways for biosynthesis of polysaccharides and peptidopolysaccharides. Collectively, secondary metabolism involves all primary pathways of carbon metabolism (Fungal Physiology, Chapter 9 pp 246-274 ed DH Griffin (1994)). "Secondary metabolites" also include intermediate compounds in the biosynthetic pathway for a secondary metabolite that are dedicated to the pathway for synthesis of the secondary metabolite. "Dedicated to the pathway for synthesis of the secondary metabolite" means that once the intermediate is synthesized by the cell, the cell will not convert the intermediate to a primary metabolite. "Intermediate compounds" also include secondary metabolite intermediate compounds which can be converted to useful compounds by subsequent chemical conversion or subsequent biotransformation. As such, providing improved availability of such intermediate compounds would still lead to improved production of the ultimate useful compound, which itself may be referred to herein as a secondary metabolite. The yeast *Saccharomyces cerevisiae* is not known to produce secondary metabolites. The term "primary metabolite" means a natural product that has an obvious role in the functioning of almost all organisms. Primary metabolites include, without limitation, compounds involved in the biosynthesis of lipids, carbohydrates, proteins, and nucleic acids. The term "increasing the yield of the secondary metabolite" means increasing the quantity of the secondary metabolite present in the total fermentation broth per unit volume of fermentation broth or unit mass of solid media. An "extracellular enzyme" is an enzyme that is secreted by a fungus into its surrounding media.

The following examples are intended to further illustrate certain particularly preferred embodiments of the invention and are not intended to limit the scope of the invention.

Example 1

Transformation of filamentous fungi

Protoplasts were generated as follows. Rich media was inoculated with spores and spores were allowed to germinate for about 20 hrs or until germ tubes were between 5 and 10 spore lengths. If a fungus does not sporulate (some taxol producing fungi do not sporulate), hyphae can be sonicated into small fragments and these small hyphal fragments can be used as inoculums. Another alternative is to find a medium where the fungus will sporulate. For example, sporulation of a taxol producing fungus can be induced by growth on water agar plates that contain gamma-irradiated carnations leaves. In addition, high salt medium can promote sporulation in late sporulating *P. chrysogenum* strains. The germlings were centrifuged and washed twice with sterile distilled water and once with 1M magnesium sulfate. Germlings were then resuspended in 1M magnesium sulfate containing approximately 2mg/ml of Novozyme. Tubes were then incubated at 30°C shaking at 80 rpm for about 2 hrs or until most of the hyphae were digested and protoplasts were abundant. Protoplasts were filtered through one layer of Miracloth. At least one volume of STC was added and protoplasts were spun down. Protoplasts were washed twice with STC. Protoplasts then were resuspended in 1ml STC and counted in a hemocytometer. A final concentration of approximately 5×10^7 protoplasts/ml were frozen in a 9:1:0.1 solution of STC, SPTC and DMSO in a Nalgene Cryo cooler at -80°C (cools -1°C/min).

Solutions for transformation were as follows: STC (0.8M Sorbitol, 25mM Tris-HCl pH 7.5, 25mM $CaCl_2$) and SPTC (0.8M Sorbitol, 40% PEG 4000, 25mM Tris-HCl pH 8, 50mM $CaCl_2$).

1-5µg of DNA comprising a novel regulator gene according to the invention in a fungal expression vector was placed in a 50 mL Falcon tube. 100µl of previously frozen protoplasts were added to the DNA, gently mixed, and then incubated on ice for 30 min. 15µl of SPTC was added, followed by mixing by tapping and incubation at RT for 15 min. 500µl SPTC was added and mixed well by tapping and rolling, then incubated at RT

for 15 min. 25 mls of regeneration minimal medium was added, mixed well and poured on plates containing 25 mls of regeneration minimal medium with 2X the concentration of selection drug. Transformation plates were incubated at 26°C for 5-6 days or until colonies started to appear. Regeneration minimal medium contains trace elements, salts, 25mM sodium nitrate, 0.8M Sucrose, and 1% agarose at pH 6.5. The selection drug that was used successfully with *P. chrysogenum* and *A. terreus* is phleomycin, a broad-spectrum glycopeptide antibiotic. Transformants were picked onto new plates with a toothpick (if fungus was sporulating) or with sterile forceps (if fungus did not sporulate). Purification plates contained minimal medium (same as regeneration minimal medium but containing 2% instead of 0.8M sucrose) and 1X drug concentration. Picked transformants were incubated at 26°C for 5-6 days.

Example 2

Growing transformants in production media

For *P. chrysogenum* and penicillin production, a plug containing spores and mycelia was used as the inoculum. The medium used was the published P2 production media which contains, 30% lactose, 5X pharmamedia cotton seed flour, ammonium sulfate, calcium carbonate, potassium phosphate, potassium sulfate, and phenoxyacetic acid pH 7. Flasks were incubated at 26°C with shaking at 225 rpm.

For *A. terreus* and lovastatin production, spores were used as the inoculum. Spores were obtained from the purification plate by using a wooden inoculation stick. The medium was RPM containing corn steep liquor, sodium nitrate, potassium phosphate, magnesium sulfate, sodium chloride, P2000 (Dow chemical), trace elements and lactose or glucose as carbon source. The medium was pH 6.5. Flasks were incubated at 26°C with shaking at 225 rpm. For static 96-well cultures, the same medium was used and the spores were obtained from the purification plate with a wooden toothpick. 96-well plates were incubated, without shaking at 26°C.

Sampling was done after 6 days of incubation for penicillin, after 5 days for lovastatin, and after 21 days for taxol. For shake flask experiments 1-1.5 mls of supernatant was placed into 96-well plates, which were centrifuged and supernatants

transferred to new 96-well plates. Samples were frozen at $^{-}80^{o}C$ for storage or for later assays.

Cultures that were grown standing in a 96-well plate were centrifuged and the supernatant was transferred to a new 96 well plate. Samples were frozen at $^{-}80^{o}C$.

The assay kit from Hawaii Biotechnology Group has a taxol detection limit of 0.5 ng/mL. Assuming a fungus which produces taxol at a level of 25ng/L, a 20-25-fold concentration step of the taxol-producing culture should be within the standard curve achievable with the kit (>0.625ng/mL). In order to have enough of each sample for duplicate testing, we will need a minimum culture volume of 15 mL. Samples will be taken after 21 days of incubation.

Example 3
Determination of penicillin concentration

Solutions of phenoxymethylpenicillin (sodium salt) in 10 mM potassium phosphate (pH 7.0) were prepared at 0, 25, 50, 100, 200, 300, 400 and 500 μg/mL. Imidazole reagent was prepared as follows. 8.25 g of imidazole was dissolved in 60 mL of water, 10 mL of 5 M HCl was added and then 10 mL of a solution of mercuric chloride (0.27 g dissolved in 100 mL of water) was added. The pH was adjusted to 6.80 +/- 0.05 with 5 M HCl and the solution was diluted to 100 mL with water.

The fermentation broth was clarified by centrifugation for 10 min at 4000 g. 40 μL of clarified fermentation broth and penicillin standard solutions were pipetted into individual wells of a 96-well UV, collection plate. 200 μL of imidazole reagent was pipetted into a 96-well filter plate (0.45 micron). The derivatization reaction of penicillin was initiated by vacuum filtration of imidazole reagent into a collection plate containing the aliquoted samples and standards. The collection plate was placed into the 96-well plate reader at 45 ^{o}C while an increase at 325 nm was monitored over 20 minutes.
A Molecular Devices 96-well UV/Vis plate reader was used for all spectrophotometric detection. This is a conventional assay, described, *e.g.*, in Bundgaard and Ilver, Journal of Pharm Pharmac 24: 790-794 (1972).

Example 4

Determination of lovastatin concentration

10 μL of sample was removed and diluted 1:100 in H_2O. 10 μl of this diluted broth was assayed in a reaction (200 μL total) containing 1 mM HMGCoA, 1 mM NADPH, 0.005 mM DTT and 5 μL $(His)_6$HMGR. The disappearance of absorbance at 340 nm was observed over time. This represents the disappearance of NADPH, and lovastatin inhibits this reaction. The initial velocities were calculated for the reactions containing samples, adjusted for dilution, and compared to reactions containing lovastatin standards to determine levels of metabolite produced. $(His)_6$HMGR was expressed in *Saccharomyces cerevisiae* and purified with a nickel column

Example 5

Determination of taxol concentration

Assay 1: From a resting culture of at least 15 mL, the broth will be neutralized to pH 7 (taxol is unstable in non-neutral conditions), frozen using liquid nitrogen and the entire sample lyophilized to dryness (cell mass and broth together) overnight. The dry material will be pulverized using glass beads and vigorous shaking. 10 mL of methylene chloride will be used to extract each sample (estimated at 1 g of dry material). After 6 hours, the methylene chloride extract will be passed through a cotton plug and lyophilized to dryness overnight. 50 μL of methanol will be added to every sample and once in solution, 450 μL of appropriate buffer or water will be added and tested by immunoassay. If we assume that there was 15 mL of starting sample from a fungus that produced approximately 25ng/L, then the final concentration is expected to be approximately 0.75 ng/mL.

Assay 2: From a resting culture of at least 15 mL, the broth will be neutralized to pH 7 and filtered through four layers of cheesecloth. The broth will be extracted with 15 mL of methylene chloride, twice and the organic layers combined. The biomass will be macerated and extracted three times with 10 mL each of chloroform/methanol (1:1). The extracts of macerated mycelia will be combined with the extracted broths, and the 60 mL

of organic extract per sample will be lyophilized to dryness overnight . 50 μL of methanol will be added to every sample and once in solution, 450 μL of appropriate buffer or water will be added and tested by immunoassay. If we assume that there were 15 mLs of starting sample from a fungus that produced approximately 25ng/L, then the final concentration is expected to be approximately 0.75 ng/mL.

The assay kit includes a taxol-protein antigen that is detected by an anti-taxane antibody. The presence of taxol in a tested isolate competitively inhibits detection of the taxol-protein antigen. The immunoassay requires 100 μL samples and follows the protocol in the kit. It will be done in a 96-well plate and will take approximately 8-9 hours.